



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 18, 2017 – 01:32 PM EST

PDB ID : 5X5I
Title : The X-ray crystal structure of a TetR family transcription regulator RcdA involved in the regulation of biofilm formation in Escherichia coli
Authors : Sugino, H.; Usui, M.; Shimada, T.; Nakano, M.; Ogasawara, H.; Ishihama, A.; Hirata, A.
Deposited on : 2017-02-16
Resolution : 2.55 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030736
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030736

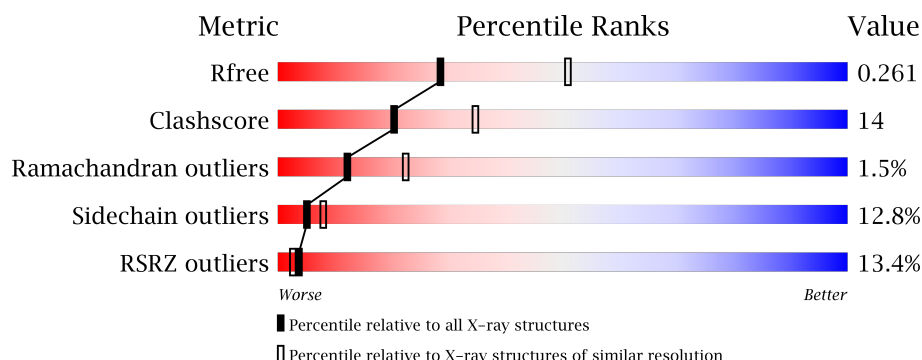
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.55 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	3689 (2.60-2.52)
Clashscore	112137	4096 (2.60-2.52)
Ramachandran outliers	110173	4037 (2.60-2.52)
Sidechain outliers	110143	4037 (2.60-2.52)
RSRZ outliers	101464	3700 (2.60-2.52)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	189	<div> <div>12%</div> <div> <div></div> <div>54%</div> <div>27%</div> <div>8%</div> <div>11%</div> </div> </div>
1	B	189	<div> <div>10%</div> <div> <div></div> <div>56%</div> <div>32%</div> <div>8%</div> </div> </div>
1	C	189	<div> <div>16%</div> <div> <div></div> <div>52%</div> <div>32%</div> <div>6%</div> <div>10%</div> </div> </div>
1	D	189	<div> <div>13%</div> <div> <div></div> <div>61%</div> <div>24%</div> <div>6%</div> <div>8%</div> </div> </div>
1	E	189	<div> <div>10%</div> <div> <div></div> <div>52%</div> <div>35%</div> <div>5%</div> <div>8%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	189	<div><div></div><div>9%</div><div>60%</div><div>25%</div><div>5%</div><div>10%</div></div>
1	G	189	<div><div></div><div>13%</div><div>58%</div><div>29%</div><div>5%</div><div>8%</div></div>
1	H	189	<div><div></div><div>12%</div><div>67%</div><div>20%</div><div>•</div><div>10%</div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10940 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called HTH-type transcriptional regulator RcdA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	169	Total	C	N	O	S	Se	0	0	0
			1349	859	229	251	1	9			
1	B	173	Total	C	N	O	S	Se	0	0	0
			1377	877	234	256	1	9			
1	C	171	Total	C	N	O	S	Se	0	0	0
			1361	867	230	254	1	9			
1	D	173	Total	C	N	O	S	Se	0	0	0
			1367	869	230	258	1	9			
1	E	174	Total	C	N	O	S	Se	0	0	0
			1378	876	232	260	1	9			
1	F	171	Total	C	N	O	S	Se	0	0	0
			1362	867	231	254	1	9			
1	G	174	Total	C	N	O	S	Se	0	0	0
			1380	878	235	257	1	9			
1	H	171	Total	C	N	O	S	Se	0	0	0
			1356	865	229	252	1	9			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	179	ALA	-	expression tag	UNP P75811
A	180	ALA	-	expression tag	UNP P75811
A	181	ALA	-	expression tag	UNP P75811
A	182	LEU	-	expression tag	UNP P75811
A	183	GLU	-	expression tag	UNP P75811
A	184	HIS	-	expression tag	UNP P75811
A	185	HIS	-	expression tag	UNP P75811
A	186	HIS	-	expression tag	UNP P75811
A	187	HIS	-	expression tag	UNP P75811
A	188	HIS	-	expression tag	UNP P75811
A	189	HIS	-	expression tag	UNP P75811
B	179	ALA	-	expression tag	UNP P75811
B	180	ALA	-	expression tag	UNP P75811

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	181	ALA	-	expression tag	UNP P75811
B	182	LEU	-	expression tag	UNP P75811
B	183	GLU	-	expression tag	UNP P75811
B	184	HIS	-	expression tag	UNP P75811
B	185	HIS	-	expression tag	UNP P75811
B	186	HIS	-	expression tag	UNP P75811
B	187	HIS	-	expression tag	UNP P75811
B	188	HIS	-	expression tag	UNP P75811
B	189	HIS	-	expression tag	UNP P75811
C	179	ALA	-	expression tag	UNP P75811
C	180	ALA	-	expression tag	UNP P75811
C	181	ALA	-	expression tag	UNP P75811
C	182	LEU	-	expression tag	UNP P75811
C	183	GLU	-	expression tag	UNP P75811
C	184	HIS	-	expression tag	UNP P75811
C	185	HIS	-	expression tag	UNP P75811
C	186	HIS	-	expression tag	UNP P75811
C	187	HIS	-	expression tag	UNP P75811
C	188	HIS	-	expression tag	UNP P75811
C	189	HIS	-	expression tag	UNP P75811
D	179	ALA	-	expression tag	UNP P75811
D	180	ALA	-	expression tag	UNP P75811
D	181	ALA	-	expression tag	UNP P75811
D	182	LEU	-	expression tag	UNP P75811
D	183	GLU	-	expression tag	UNP P75811
D	184	HIS	-	expression tag	UNP P75811
D	185	HIS	-	expression tag	UNP P75811
D	186	HIS	-	expression tag	UNP P75811
D	187	HIS	-	expression tag	UNP P75811
D	188	HIS	-	expression tag	UNP P75811
D	189	HIS	-	expression tag	UNP P75811
E	179	ALA	-	expression tag	UNP P75811
E	180	ALA	-	expression tag	UNP P75811
E	181	ALA	-	expression tag	UNP P75811
E	182	LEU	-	expression tag	UNP P75811
E	183	GLU	-	expression tag	UNP P75811
E	184	HIS	-	expression tag	UNP P75811
E	185	HIS	-	expression tag	UNP P75811
E	186	HIS	-	expression tag	UNP P75811
E	187	HIS	-	expression tag	UNP P75811
E	188	HIS	-	expression tag	UNP P75811
E	189	HIS	-	expression tag	UNP P75811

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	179	ALA	-	expression tag	UNP P75811
F	180	ALA	-	expression tag	UNP P75811
F	181	ALA	-	expression tag	UNP P75811
F	182	LEU	-	expression tag	UNP P75811
F	183	GLU	-	expression tag	UNP P75811
F	184	HIS	-	expression tag	UNP P75811
F	185	HIS	-	expression tag	UNP P75811
F	186	HIS	-	expression tag	UNP P75811
F	187	HIS	-	expression tag	UNP P75811
F	188	HIS	-	expression tag	UNP P75811
F	189	HIS	-	expression tag	UNP P75811
G	179	ALA	-	expression tag	UNP P75811
G	180	ALA	-	expression tag	UNP P75811
G	181	ALA	-	expression tag	UNP P75811
G	182	LEU	-	expression tag	UNP P75811
G	183	GLU	-	expression tag	UNP P75811
G	184	HIS	-	expression tag	UNP P75811
G	185	HIS	-	expression tag	UNP P75811
G	186	HIS	-	expression tag	UNP P75811
G	187	HIS	-	expression tag	UNP P75811
G	188	HIS	-	expression tag	UNP P75811
G	189	HIS	-	expression tag	UNP P75811
H	179	ALA	-	expression tag	UNP P75811
H	180	ALA	-	expression tag	UNP P75811
H	181	ALA	-	expression tag	UNP P75811
H	182	LEU	-	expression tag	UNP P75811
H	183	GLU	-	expression tag	UNP P75811
H	184	HIS	-	expression tag	UNP P75811
H	185	HIS	-	expression tag	UNP P75811
H	186	HIS	-	expression tag	UNP P75811
H	187	HIS	-	expression tag	UNP P75811
H	188	HIS	-	expression tag	UNP P75811
H	189	HIS	-	expression tag	UNP P75811

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	3	Total O 3 3	0	0
2	D	1	Total O 1 1	0	0
2	E	1	Total O 1 1	0	0

Continued on next page...

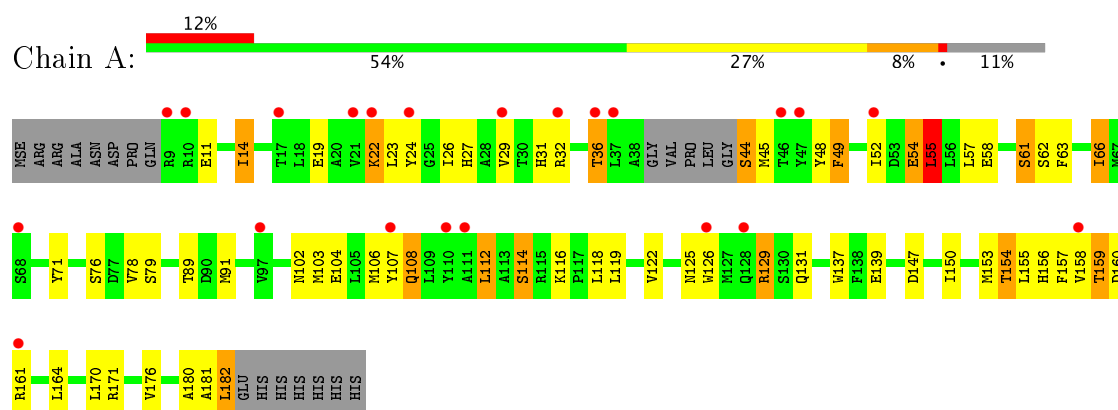
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	2	Total	O	0	0
			2	2		
2	H	3	Total	O	0	0
			3	3		

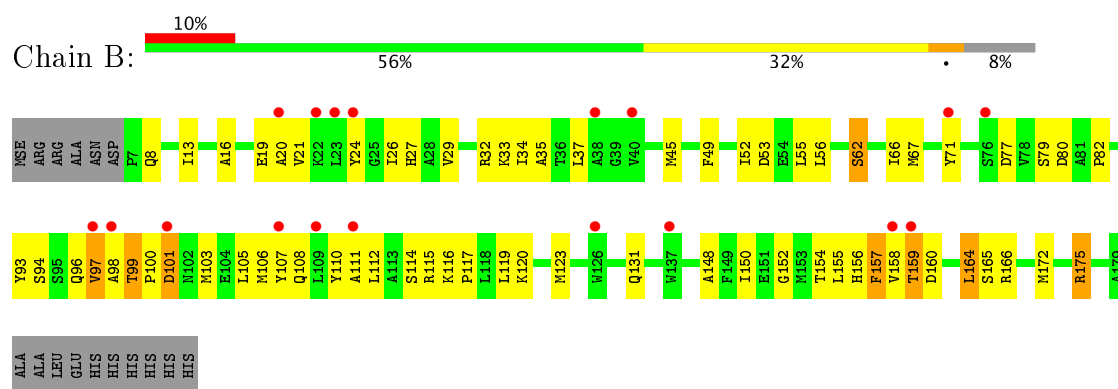
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

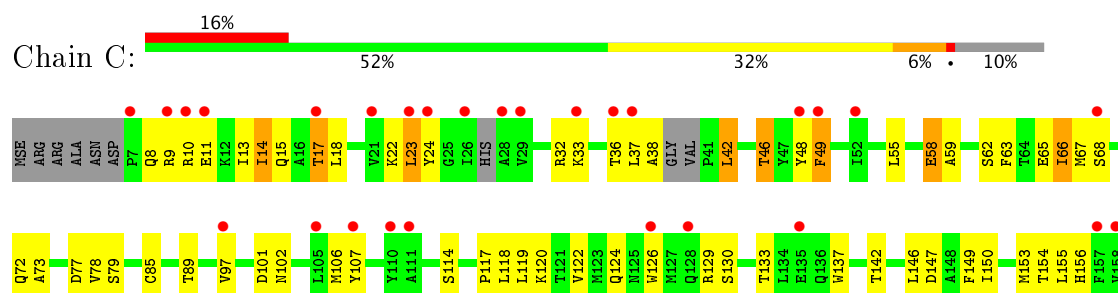
- Molecule 1: HTH-type transcriptional regulator RcdA

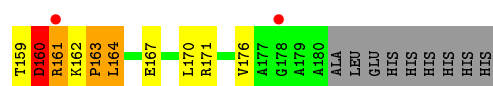


- Molecule 1: HTH-type transcriptional regulator RcdA



- Molecule 1: HTH-type transcriptional regulator RcdA

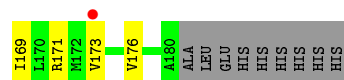
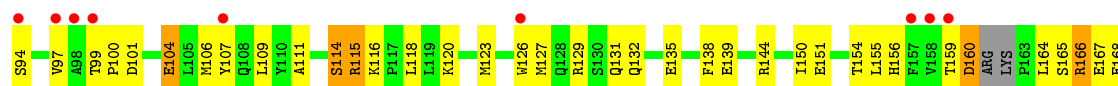
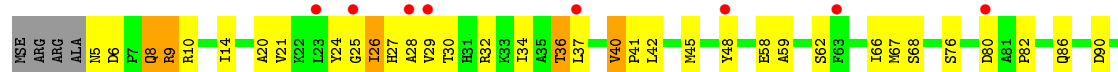




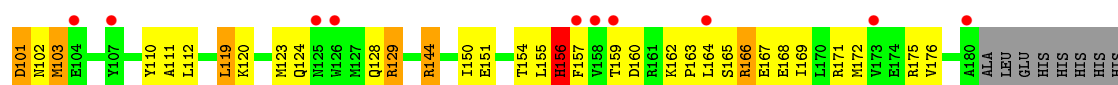
- Molecule 1: HTH-type transcriptional regulator RcdA



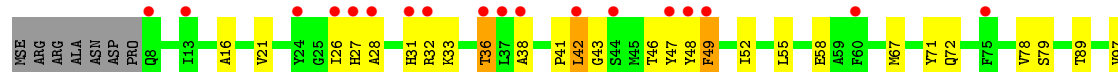
- Molecule 1: HTH-type transcriptional regulator RcdA

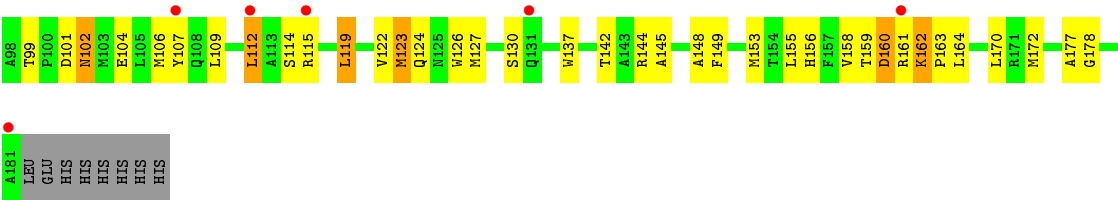


- Molecule 1: HTH-type transcriptional regulator RcdA

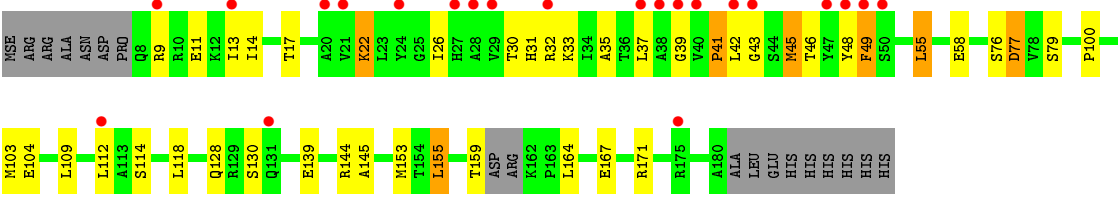


- Molecule 1: HTH-type transcriptional regulator RcdA





● Molecule 1: HTH-type transcriptional regulator RcdA



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	38.48Å 75.97Å 110.05Å 90.00° 90.01° 89.95°	Depositor
Resolution (Å)	37.98 – 2.55 37.98 – 2.55	Depositor EDS
% Data completeness (in resolution range)	98.3 (37.98-2.55) 96.4 (37.98-2.55)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	6.54 (at 2.54Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.198 , 0.267 0.192 , 0.261	Depositor DCC
R_{free} test set	1939 reflections (4.94%)	DCC
Wilson B-factor (Å ²)	44.4	Xtriage
Anisotropy	0.256	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 33.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.430 for h,-k,-l 0.429 for -h,k,-l 0.487 for -h,-k,l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	10940	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.15% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z > 5$	RMSZ	# $ Z > 5$
1	A	0.64	2/1365 (0.1%)	0.83	3/1828 (0.2%)
1	B	0.51	0/1396	0.71	2/1872 (0.1%)
1	C	0.52	0/1377	0.78	2/1842 (0.1%)
1	D	0.52	0/1385	0.69	0/1858
1	E	0.51	0/1396	0.69	1/1873 (0.1%)
1	F	0.51	0/1380	0.69	0/1848
1	G	0.50	0/1398	0.73	2/1875 (0.1%)
1	H	0.54	0/1373	0.73	0/1840
All	All	0.53	2/11070 (0.0%)	0.73	10/14836 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	2
1	D	0	1
1	F	0	1
1	G	0	2
1	H	0	1
All	All	0	7

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	49	PHE	CE1-CZ	-8.73	1.20	1.37
1	A	49	PHE	CG-CD2	-6.91	1.28	1.38

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	54	GLU	C-N-CA	8.03	141.77	121.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	164	LEU	CA-CB-CG	-7.04	99.11	115.30
1	G	178	GLY	N-CA-C	6.39	129.07	113.10
1	A	55	LEU	CA-CB-CG	-5.65	102.31	115.30
1	C	163	PRO	C-N-CA	5.51	135.48	121.70

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	160	ASP	Peptide
1	C	161	ARG	Peptide
1	D	96	GLN	Peptide
1	F	164	LEU	Peptide
1	G	162	LYS	Peptide

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1349	0	1347	48	0
1	B	1377	0	1376	44	0
1	C	1361	0	1361	50	0
1	D	1367	0	1349	45	0
1	E	1378	0	1364	55	0
1	F	1362	0	1355	42	1
1	G	1380	0	1378	35	1
1	H	1356	0	1355	32	0
2	A	3	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	G	2	0	0	0	0
2	H	3	0	0	1	0
All	All	10940	0	10885	306	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 14.

The worst 5 of 306 close contacts within the same asymmetric unit are listed below, sorted by

their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:72:GLN:NE2	1:C:133:THR:OG1	1.97	0.98
1:C:10:ARG:NH1	1:C:58:GLU:OE1	1.98	0.96
1:E:99:THR:HG22	1:E:100:PRO:HD2	1.55	0.89
1:H:41:PRO:HB2	1:H:42:LEU:HA	1.56	0.84
1:A:26:ILE:HG23	1:A:27:HIS:H	1.44	0.83

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:44:SER:OG	1:G:97:VAL:O[1_665]	2.17	0.03

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	165/189 (87%)	152 (92%)	11 (7%)	2 (1%)	15	29
1	B	171/189 (90%)	160 (94%)	7 (4%)	4 (2%)	7	12
1	C	165/189 (87%)	152 (92%)	12 (7%)	1 (1%)	28	49
1	D	169/189 (89%)	153 (90%)	14 (8%)	2 (1%)	15	29
1	E	170/189 (90%)	159 (94%)	11 (6%)	0	100	100
1	F	167/189 (88%)	155 (93%)	9 (5%)	3 (2%)	10	17
1	G	172/189 (91%)	157 (91%)	9 (5%)	6 (4%)	4	5
1	H	167/189 (88%)	158 (95%)	7 (4%)	2 (1%)	15	29
All	All	1346/1512 (89%)	1246 (93%)	80 (6%)	20 (2%)	12	22

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	55	LEU
1	D	26	ILE
1	F	163	PRO
1	G	28	ALA
1	G	160	ASP

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	143/150 (95%)	121 (85%)	22 (15%)	3	4
1	B	147/150 (98%)	130 (88%)	17 (12%)	6	10
1	C	145/150 (97%)	125 (86%)	20 (14%)	4	6
1	D	145/150 (97%)	128 (88%)	17 (12%)	6	10
1	E	147/150 (98%)	128 (87%)	19 (13%)	5	8
1	F	145/150 (97%)	127 (88%)	18 (12%)	5	9
1	G	146/150 (97%)	127 (87%)	19 (13%)	5	7
1	H	144/150 (96%)	127 (88%)	17 (12%)	6	10
All	All	1162/1200 (97%)	1013 (87%)	149 (13%)	5	8

5 of 149 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	77	ASP
1	E	94	SER
1	H	55	LEU
1	D	99	THR
1	D	160	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	72	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	96	GLN
1	G	31	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	160/189 (84%)	0.87	22 (13%) 3 2	21, 52, 87, 108	0
1	B	164/189 (86%)	0.87	18 (10%) 6 5	29, 49, 100, 131	0
1	C	162/189 (85%)	1.00	30 (18%) 1 1	21, 52, 101, 128	0
1	D	164/189 (86%)	0.90	24 (14%) 3 2	27, 50, 93, 140	0
1	E	165/189 (87%)	0.80	18 (10%) 6 5	28, 49, 87, 127	0
1	F	162/189 (85%)	0.81	17 (10%) 7 6	28, 49, 90, 123	0
1	G	165/189 (87%)	0.93	24 (14%) 3 2	21, 53, 110, 133	0
1	H	162/189 (85%)	0.93	22 (13%) 3 2	21, 50, 104, 134	0
All	All	1304/1512 (86%)	0.89	175 (13%) 4 2	21, 51, 100, 140	0

The worst 5 of 175 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	38	ALA	8.7
1	H	28	ALA	8.5
1	D	158	VAL	8.1
1	B	158	VAL	7.9
1	C	36	THR	7.8

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.